NOTES 14: P-VALUES AND SIGNIFICANCE

Stat 120 | Fall 2025

Prof Amanda Luby

Last time, we laid out the following steps for setting up a hypothesis test:

1. Formulate hypotheses in terms of *population parameter*
2. Collect data and compute a *sample statistic*
3. Use the *sample statistic* to make a claim about the *hypotheses*

Today, we're diving into step 3:

3a. Construct *null/randomization distribution*

If H_0 is _____, what does the range of sample statistics look like?

3b. Compute *p-value*

Numeric summary of how extreme _____
is compared to _____ distribution

3c. Make a formal decision about H_0

- Reject H_0 if _____
- Do not Reject H_0 _____

StatKey Example 1 In a random sample of 765 adults in the United States, 322 say they could not cover a $400 unexpected expense without borrowing money or going into debt. A journalist claims that the true proportion is actually 50%. Is the journalist justified?

Constructing null distribution for a single proportion

In this example, flip a coin 765 times and record the proportion that came up heads. Do this many times to construct the randomization distribution

StatKey Example 2 In a Pew Research Poll on social media use, 72% of Twitter users (249 out of 346) responded that they visited Twitter a few times a week or more. Among Instagram users, this number was 80% (424 out of 530). Is there a difference in frequency of use between Twitter and Instagram users?

Constructing null distribution for a difference in proportions

Combine all cases, then randomly assign each case to one of two groups with the same sample sizes as the original data. Record the difference in proportions between the two groups. Do this many times.

R Example 1

Here is what the dataset from the last example might look like in R:

```
# A tibble: 876 × 2
   app       few_times_or_more
   <chr>                 <dbl>
 1 instagram                 1
 2 twitter                   1
 3 twitter                   0
 4 instagram                 1
```

```
library(CarletonStats)
permTest(few_times_or_more ~ app, data = social_media, seed = 101525)
```

Check:

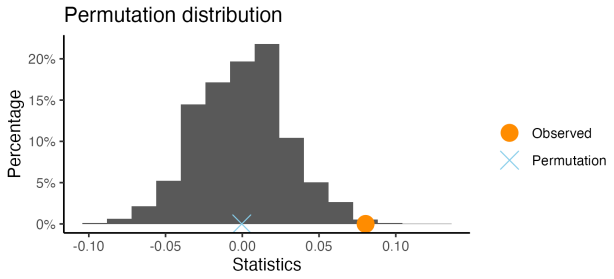
This is a different p-value than StatKey - why?

```
   ** Permutation test **

Permutation test with alternative: two.sided
Observed statistic
 instagram :  0.8   twitter :  0.7197
Observed difference: 0.08035

Mean of permutation distribution: -0.00034
Standard error of permutation distribution: (
P-value:  0.0057

   *-------------*
```



R Example 2 In a study about Universal Basic Income, the (SEED) randomly selected 100 residents and gave them an unrestricted $500 a month for 24 months. At the end of the study period, they were given an emotional well-being assessment and results were compared to 100 control-group residents who did not receive payments. The average well-being score among the UBI group was 31.5 and among the control group was 29.5. Does UBI result in higher well-being scores?

```
# A tibble: 2 x 3
  group    mean     n
  <chr>   <dbl> <int>
1 Control  29.5   100
2 UBI      31.5   100
```

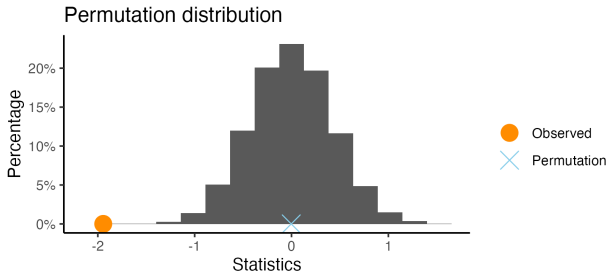
```r
permTest(well_being ~ group, data = ubi, alternative = "greater", seed = 101525)
```

```
   ** Permutation test **

Permutation test with alternative: greater
Observed statistic
 Control :  29.54   UBI :  31.49
Observed difference: -1.944

Mean of permutation distribution: -0.00282
Standard error of permutation distribution: (
P-value:  1

   *-------------*
```



Constructing null distribution for a difference means

Combine all cases, then randomly assign each case to one of two groups with the same sample sizes as the original data. Record the difference in means between the two groups. Do this many times.

Statistical Significance

Significance Level

Cutoff value for rejecting H_0 (called "alpha" or α)
If p-value _____ α, we "reject H_0" and say results are "statistically significant"
Common values:
Default:

Statistical Significance

If p-value is small enough, it's very unlikely to get the results we observed if H_0 is true due to random chance.
Results are statistically significant if they provide convincing evidence against H_0

Formal Statistical Decisions

- p-value ____ α: Reject H_0
- p-value ____ α: Do not reject H_0

Summary

Whenever performing a hypothesis test, you should:

(1) Perform an appropriate EDA
(2) Formulate H_0 and H_A
(3) Compute a *test statistic*
(4) Construct a *null distribution* using `permTest` in R or StatKey
(5) Compare the test statistic to the null distribution and compute a *p-value*
(6) Make a decision (reject or fail to reject) about H_0
(7) Report your conclusion in context